

Archived version from NCDOCKS Institutional Repository <http://libres.uncg.edu/ir/asu/>

Cummings, R. G., Martinez-Vazquez, J., McKee, M., & Torgler, B. (2009). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal Of Economic Behavior & Organization*, 70(3), 447-457. doi:10.1016/j.jebo.2008.02.010

Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment

Ronald G. Cummings, Jorge Martinez-Vazquez, **Michael McKee**, Benno Torgler

Abstract

Our working hypothesis is that cross-cultural differences in tax compliance behavior have foundations in the institutions of tax administration and citizen assessment of the quality of governance. Tax compliance being a complex behavioral issue, its investigation requires use of a variety of methods and data sources. Results from artefactual field experiments conducted in countries with substantially different political histories and records of governance quality demonstrate that observed differences in tax compliance levels persist over alternative levels of enforcement. The experimental results are shown to be robust by replicating them for the same countries using survey response measures of tax compliance.

Keywords:

Tax compliance, Governance, Artefactual field experiments

1. Introduction

One of the more vexing problems for policy makers in developing and transition economies is encouraging high levels of tax compliance. This issue is independent of the overall tax “take” from GDP. As Cowell (1990, p. 6) notes, “the issue of evasion is, unlike other illegal activities, inseparably bound up with the instruments of fiscal control that the government attempts to use in carrying out its economic policy.” While reducing evasion improves the government’s revenue, it is a broader issue for the development of a civil order (Knack and Keefer, 1997).

However, reducing tax evasion is often not simply a matter of applying higher penalties and/or increasing the frequency of audits. Extreme penalties may backfire by creating a setting in which bribery and corruption are more prevalent with the end result being lower tax compliance and a general loss of trust in public institutions. Designing effective policies for reducing tax evasion requires understanding the *behavioral* aspects of the tax compliance decision. Individual attitudes toward compliance have been shown to be a function of social and cultural norms (Elster, 1989; Naylor, 1989). If such attitudes would carry over to actual compliance, enhancing these norms, as through increasing overall trust in government, is a desirable policy instrument to complement the usual enforcement options. Botswana and South Africa provide a good test of this hypothesis. These countries have experienced strikingly different social histories, despite being geographic neighbors and achieving independence at about the same time. Thus, these countries offer a natural, or field, experiment for the investigation of the effects of tax morale stemming from perceptions of public institutions.

We report on analyses of data derived from surveys and from an artefactual field experiment to investigate the effects of political norms on compliance behavior. Tax compliance is a complex behavioral issue, and investigation requires the use of a variety of methods and data sources as each instrument has strength and weaknesses. The use of survey and experimental data permits a broader exploration of the effects of such norms on tax compliance behavior and provides a robustness check. Our approach is to conduct an artefactual field experiment (following the taxonomy presented by Harrison and List, 2004) utilizing the comparative advantage of laboratory experiments (the potential to control for extraneous factors and to manipulate the variables of interest) in conjunction with the richer decision setting and participant pool that is available in the field. It is this richer setting and a participant pool consisting largely of working adults that allows political norms into our analysis and a test of the effect of these on compliance behavior.

The experimental and survey results reported in this paper provide support for the hypothesis that tax compliance increases with individual perceptions of good governance. The responses to the usual enforcement mechanisms (audits and penalties) are shown to be enhanced by these perceptions, which we can call tax morale, defined as the *intrinsic motivation to pay taxes* arising from the moral obligation to pay taxes as a contribution to society. In the settings investigated via our artefactual field experiments, baseline compliance varies as expected given the political and social

histories of the two countries, and while compliance does increase with enforcement effort, the effect is less in the country for which governance is less good. The results reported provide support for a model of tax compliance behavior that extends beyond the typical “economics of crime” approach with its emphasis on enforcement effort and deterrence.

Analyses of data from surveys of public attitudes toward government show that perceptions of fairness and efficacy are considerably higher in Botswana and *self-reported* tax compliance is higher. Tax compliance is difficult to observe in the field since it is an illegal, hence hidden, activity and direct observations are available only for the small fraction of taxpayers actually audited. The results from our experiments, which allow us to implement identical enforcement policies in both countries, demonstrate differences in compliance behavior consistent with the social norms implied from the survey data. Similarities in the tax systems permit use of the same jargon in each country and since English is widely used in both countries we are able to conduct the experiments in a common language. These features facilitate our comparison of tax compliance behavior across these otherwise divergent cultures. Our artefactual field experiments were conducted in the fall of 1999 and the survey data were collected during 1999–2000.

Deficiencies in field data are well known and true field experiments are rare owing to the costs and complexities of manipulating the actual tax collection system. A creative attempt in the use of a field experiment is Slemrod et al. (2001) who, working with the State of Minnesota, sent a letter (in January 1994) to a stratified random sample of roughly 2000 Minnesota taxpayers informing them that the tax returns they were about to file would be “closely examined”. This experiment was designed to learn whether informing individuals about an increase in the probability of audit prior to filing a tax return would in fact increase their compliance. In the absence of actual audits, the focus was on investigating the income reporting behavior of the taxpayers, an indirect measure of compliance.

2. The analytics of the tax compliance decision

The basic compliance model is based on Allingham and Sandmo (1972) and Yitzhaki (1974). Suppose that an individual receives a fixed amount of income I and must choose how much to declare to the tax authorities. Declared income D is taxed at the rate t . Unreported income is not taxed; however, the individual may be audited with probability p , at which point a fine f is imposed on each dollar of unpaid taxes. For the interesting case where $D < I$, if underreporting is detected the individual's income IC equals $IC = I - tD - ft(I - D)$, while, if underreporting is not detected income IN is $IN = I - tD$. The individual chooses D to maximize the expected utility $EU(I)$ of the evasion gamble, or $EU(I) = pU(IC) + (1-p)U(IN)$, where utility $U(I)$ is assumed to be a function only of income. This optimization generates the first-order condition, $pU'(IC)(f-1)t - (1-p)U'(IN)t = 0$. This is the basic portfolio model of tax compliance. It is straightforward to show, within this model, that increases in the probability of an audit and/or the fine rate will increase compliance.

Given the enforcement resources available to most governments, the observed high compliance rates are inconsistent with rational behavior. Uncertainty regarding the actual audit practices may play a role. Audit probabilities are largely subjective since the tax authority does not have an incentive to reveal the entire audit mechanism (Alm, 1988) and individuals may

Table 1 Features of the tax system in the study countries (effective in year 2000).

Tax feature	South Africa	Botswana
Self-reporting/assessment	Yes	Yes
Withholding	Yes	Yes
Highest marginal rate	45%	25%
Audit enforcement		
Financial penalty	Yes (max: double tax owed plus interest)	Yes (max: tax owed plus interest)
Incarceration?	Yes (up to 2 years)	Yes (up to 1 year)
Mandatory filing	No (unless tax owed)	No (unless tax owed)
Central government tax amnesty	Not prior to 2000	Yes (in 1999)

have a tendency to overweight the probability of an audit. Such behavior could support high levels of compliance even with low objective probability of an audit (Bernasconi, 1998). Nevertheless, extreme degrees of risk aversion would be required to explain observed levels of compliance. Other factors must be at work.

Tax compliance is enhanced when individuals view the paying of taxes as a fair fiscal exchange. In such situations compliance is likely to increase, *ceteris paribus*. In particular, when the services provided by the government are viewed as widely desired and the decisions determining the services provided are transparent and fair, compliance is likely to be higher. This latter factor is not captured in the conventional portfolio model of tax compliance. Nevertheless, it is clear that these interactive effects may affect tax compliance decisions.

The manner by which the public expenditure budget is determined is likely to have an effect on the level of compliance. Alm et al. (1993) find that compliance is higher when the public good is voted on rather than imposed and when the political outcome is known to be widely supported. The means by which enforcement rules are determined can also influence compliance (Alm et al., 1999). Social norms and morals have been cited as reasons for high compliance with rules (Elster) and collective actions (Naylor). Even simple personal ethics may affect tax compliance behavior independently of the fiscal exchange between the government and the taxpayers (Steenbergen et al., 1992). Taken together these factors would lead us to modify the condition for optimal reporting portfolio to $pU_{-}(IC)(f-1)t - (1-p)U_{-}(IN - _)t = 0$, where $_$ denotes the psychic cost associated with evading one's own tax liability if one is not caught. The greater the moral support for government, the higher is the size of $_$ and the lower the utility from cheating. The psychic costs associated with cheating arise only if one is actually cheating. A taxpayer who complies fully and is not audited ($I=D$) experiences no change in utility. The genesis of social norms is the interesting issue (Alm and Martinez-Vazquez, 2003). If these norms evolve from perceptions that the quality of governance is high, we predict we will observe differences in tax compliance that are correlated with these perceptions. With these factors in mind, we review the basic features of the

personal income tax system and the perceptions of governance in South Africa and Botswana.

3. The tax systems and perceptions of governance

A comparison of tax morale and compliance between Botswana and South Africa constitutes a good experiment since the tax systems themselves are similar but the political histories of the two countries could scarcely be more dissimilar.

This section begins by describing features of the personal income tax (PIT) systems. The elements of the tax structure are summarized in Table 1. For the PIT, the self-assessment and audit processes are similar in both countries although there are varying degrees of aggressiveness in enforcement. Both countries rely on some form of tax withholding at source and individual self-assessment and reporting of final tax liabilities. Tax evasion is treated as a serious crime in South Africa; the tax authority exploits high profile cases to reinforce its reputation for tough enforcement. The Botswana tax authority seems to be (relatively) more accommodating. For example, a general tax amnesty was conducted in 1999. This had not happened in South Africa when the present study was conducted.

The Botswana investigative division carries out audits in cases where tax evasion is suspected. Civil penalties can also be imposed for failure to file if taxes are owed. These penalties consist of interest (two percent per month) and a penalty up to the tax owed. Criminal penalties up to 1 year can be imposed for egregious evasion and/or fraud. The South Africa penalty structure is generally harsher; a person required to file a return who fails to do so within the period mentioned above is liable to a penalty up to R2000 and/or to imprisonment for a period up to 12 months. Further, taxable income may be estimated and three times the amount of tax charged thereon. Any taxpayer who knowingly and willfully makes false statements in the return or evades or attempts to evade taxation (and any person who assists a taxpayer to do so) is liable to a penalty up to R1000 and/or to imprisonment for a period up to 2 years. The taxpayer is, in addition, liable to be assessed and charged three times the amount of the evaded tax. The level of sophistication of the tax enforcement apparatus differs considerably between Botswana and South Africa.

The respective computations of the PIT bases are similar. In South Africa, the PIT base consists of wages and salaries as well as passive income (e.g., interest and dividends) but not capital gains. In Botswana, the PIT base includes wages and salaries as well as all investment income (interest, dividends, and capital gains). The marginal rate is capped at 25 percent in Botswana, lower than the top rate in South Africa (45 percent).

Botswana, lower than the top rate in South Africa (45 percent). It is expected that tax compliance will be affected by enforcement effort but also by the inhibitors that are inherent in the individual perception of the quality of governance. Both countries take steps to associate taxation with the provision of government services. Botswana's political history is virtually unique among African countries. Although a former colony (British) and only recently (1966) gained independence, diamond-rich Botswana is one of Africa's oldest multiparty democracies having successfully made the transition to self-

governance. Several elections have been held since independence, and all have been quiet affairs with none of the violence or corruption charges that have accompanied elections in neighboring countries. In fact, the government of Botswana takes great pride in its stability and refers to itself as the “gem of Africa” in many official publications. Acemoglu et al. (2002) report that pre-colonial tribal institutions developed by the Tswana tribes encouraged cooperation and participation and helped to constrain political elites. The Botswana experience is in marked contrast with South Africa with its well-known history of apartheid and social discord. Indeed elections in South Africa held immediately prior to our period of analysis were controversial and accompanied by violence. Both the white and black populations have reason to be suspicious of the government. The white population has been concerned about protection of property rights (especially in the face of proposals for land reform) while the black population has little reason to trust any government. The newly formed government (led initially by Nelson Mandela) had not, as of the time of our data collection, generated a record sufficiently long to establish trust. Crime rates were very high, and there was a feeling that the social order was somewhat fragile.

The perceived quality of governance institutions affects taxpayers’ willingness to comply with taxes (Smith, 1992; Smith and Stalans, 1991). Table 2 provides detailed institutional comparisons. Transparency International’s Corruption Perception Index, which relates corruption perceptions of various countries’ government, indicates considerable differences between Botswana and South Africa: Botswana’s score is some 20 percent higher (better) than South Africa’s. These results are consistent with the Quality of Governance Index provided by Kaufmann et al. (2003). Botswana has higher control of corruption than South Africa. Similarly, the rule of law index, which measures the degree of respondents’ confidence in and compliance with the rules of society, is more than three times larger in Botswana. Consequently, the respect of citizens for the state and the institutions that govern economic and social interactions is higher in Botswana. The capacity of the government to formulate and implement sound policies effectively (represented as the government effectiveness and regulatory quality) is higher in Botswana, which also has a higher level of political stability and absence of violence. Only voice and accountability are higher in South Africa, referring to the process by which governments are selected, monitored and replaced. Overall, the values of these six governance dimensions for the periods 1998 and 2000, based on several hundred variables measuring perceptions of governance and derived from 25 different data sources, clearly indicate a higher level of institutional quality in Botswana compared to South Africa. These results are also supported by the *International Country Risk Guide* (ICRG), constructed by Stephen Knack (Knack, 1999) and the IRIS Center, University of Maryland and provided by the PRS Group, which offers an alternative set of data to the Quality of Governance Index, with special emphasis on aspects affecting private foreign investment decisions.

Table 2 also shows that political rights and the level of civil liberty are similar in both countries. To measure the variable income inequality, we use the newest available dataset, Estimated Household Income Inequality (EHII), constructed by Galbraith and Kum (2005).¹¹ The GINI coefficients indicate that income inequality is slightly greater in

Botswana. The Index of Economic Freedom clearly indicates a higher fiscal burden¹² for South Africa, but also less government interventions in South Africa. The Polity IV dataset shows, in line with previous datasets, that Botswana has more stable political institutions than South Africa.

We utilize data from the Afrobarometer, a relatively new survey measuring the social, political and economic atmosphere in more than 10 countries in Africa. This dataset allows us to incorporate the newest data covering Botswana (year 1999)

Table 2 Governance and country indicators.

	Botswana	South Africa	Year
CPI ^a	6.1 (ranking 24)	5 (ranking 34)	1999
Governance indicators ^b			
Control of corruption	0.53	0.42	1998
	1.02	0.57	2000
Rule of law	0.66	0.21	1998
	0.67	0.28	2000
Regulatory quality	0.69	0.33	1998
	0.79	0.12	2000
Government effectiveness	0.52	0.17	1998
	0.98	0.43	2000
Political stability	0.89	-0.80	1998
	0.90	-0.13	2000
Voice and accountability	0.77	0.87	1998
	0.78	1.05	2000
ICRG ^c			
Composite risk rating	81.00	66.75	January, 1999
Political risk rating	76.00	69.00	January, 1999
Economic risk rating	42	31.5	January, 1999
Law and order	4	3	January, 1999
Bureaucratic quality	2	2	January, 1999
Ethnic tensions	5	3	January, 1999
Democratic accountability	3	4	January, 1999
Corruption in government	3	3	January, 1999
External conflict	10	9	January, 1999
Government stability	11	11	January, 1999
Internal conflict	12	9	January, 1999
EHII inequality ^d	48.37	44.68	1998
Index of economic freedom ^e			
Fiscal burden	2.6	4.3	1999
Government intervention	4.5	2.5	1999
Property rights	2	3	1999
Regulation	3	2	1999
Informal market	4	4	1999
Polity IV ^f			
Institutionalized democracy	9	9	1999
Regime durability	33	5	1999
Ethnic fractionalization ^g	0.410	0.752	1997–2001

and South Africa (year 2000). The Afrobarometer has a focus on self-reported compliance with the following question to measure tax honesty:

We would like to remind you that your responses to this interview are confidential. Here is a list of actions ordinary people are taking in a political system. For each of these, please tell me whether you have engaged in this activity or not?

Avoid paying income taxes.

We have coded the variable as follows: “Yes, often”, “Yes, a few times”, and “Yes, once or twice”: 0, “No, but would do it if had the chance”: 1, “No, would never do this”: 2.

Answers of “don’t know” and missing values were not coded and were dropped from the sample. We use the survey data to conduct a multivariate analysis of tax compliance in the two countries. A dummy variable is used to control for unobserved differences across the two countries, and several variables are used to control for additional factors affecting tax morale. Given the scaled ranking information of the dependent variables, we use an ordered probit estimation, and the estimating equation is nonlinear. Since this allows direct interpretation of only the sign of the estimated coefficients, we compute the marginal effects to estimate the change in the share of taxpayers (the probability of) belonging to a specific honesty rank when the independent variable increases by one unit. We present

Table 3 Determinants of tax compliance in Botswana (1999) and South Africa (2000).

Independent variables	Ordered probit								
	EQ1			EQ2			EQ3		
	Coeff.	t-ratio	Marg.	Coeff.	t-ratio	Marg.	Coeff.	t-ratio	Marg.
(a) Socio-demogr. factors									
Female	0.092 [*]	1.79	0.023	0.084	1.523	0.022	0.029	0.478	0.007
Age	0.010 ^{***}	3.265	0.003	0.012 ^{***}	4.593	0.003	0.012 ^{***}	4.348	0.003
Education	0.004	-0.746	0.001	0.009	0.484	0.002	0.016	0.832	0.004
(b) Employment Status									
Occupation				-0.140 ^{**}	-2.039	-0.036	0.136	0.999	0.034
Employer							0.73	1.405	0.184
Miner							-0.022	-0.163	-0.005
Farmer							0.351 ^{***}	3.452	0.089
Domestic							-0.052	-0.269	-0.013
Armed services/police/SEC							0.363 ^{***}	3.235	0.092
Student							0.384	0.932	0.097
Disabled							0.105	1.296	0.027
(c) Culture									
Botswana	0.238 ^{***}	5.711	0.06	0.215 ^{***}	3.256	0.055	0.236 ^{***}	3.540	0.06
Observations	3059			2752			2752		
Prob (F-statistic)	0.000			0.000			0.000		

(see Table 3) the marginal effects only for the highest honesty rank. We observe that individuals in Botswana are more compliant than those in South Africa. The marginal effects indicate that being a resident of Botswana rather than of South Africa increases the probability of reporting the highest tax honesty by around 6 percentage points, and this result is robust across various specifications. These results support our basic hypothesis that compliance will be higher for countries with “better” governance.

4. Experimental design and hypotheses

Tax evasion is, by definition, a hidden activity. Even when there are field data from audit programs, these data typically do not cover sufficient policy changes to inform on the effects of individual policy parameters on compliance. Naturally occurring field data tell us little about those not audited, the overwhelming majority of the population. Given the non-random nature of most audit regimes, such selection issues make empirical analysis problematic even if we believe we can correct for the selection through econometric techniques. Artefactual field experiments can be used to generate data to investigate responses to changes in enforcement and impose the same levels of policy action for participants in each country.

Investigating the effects of cultural norms in the laboratory raises the question as to whether it is possible to convey these norms to the participants in the lab. Since the objective of our research is to examine the effect of cultural factors and social norms on tax compliance behavior, to induce the participants to treat our controlled setting as *if* a tax compliance decision, the experimental setting incorporates tax language and terminology to encourage the participants to incorporate social norms and cultural factors in their tax reporting decision. In the experimental setting treatments involve changing basic parameters of the tax compliance enforcement system such as audit and penalty rates. The differences across the countries can be investigated as shift effects. In sum, the tax context is *emphasized* in order that the governance effect be emphasized.

4.1. Experiment design and participant decision setting

The experiment setting replicates most of the elements of the basic structure of the personal income tax system in the study countries as described in Table 1. In the experiment, individuals receive income, pay taxes on income voluntarily reported, and face a probability of audit. If they are detected cheating, they pay a financial penalty on taxes not reported. Of course, incarceration is not a possible penalty in the experimental setting. The individual compliance decision for a given set of parameters and a given cultural baseline is expected to be a function of risk attitude. All participants participated in an initial experiment designed to elicit risk attitudes. In this experiment the participants choose either a certain payoff or a gamble for ten probabilities, ranging from 0.1 to 1.0, of the high payoff from the gamble.¹⁴ Participants select either the safe or risky option for all 10 choices, and this screen disappears until the tax experiment is completed. After the tax compliance experiment is completed the risk screen reappears with the participant's previous choices indicated. One participant rolls a 10-sided die to determine which of the choices will be used to compute a payoff. For those choosing the risky option the participant rolls a second die to determine the realized payoff. The degree of risk aversion is measured by the probability of the high payoff from the risk gamble that the individual requires in order to switch from the safe gamble. This is an early variant of the experimental design used by Holt and Laury (2002). A maintained hypothesis is that risk attitudes are the same across the cultures being investigated. This is confirmed by our results as discussed later.

Since the participant pools are exposed to identical experimental parameters, observed differences in tax compliance behavior are interpreted as deriving from differences in those institutional features affecting attitudes to government (governance quality) affecting the inhibitors.

These experiments are fully computerized. The screen image the participants interact with is a simplified tax form, and the language on the screen and in the instructions describes the setting as tax reporting decision. Participants are told they have received income and are required to disclose this income to be taxed (at a stated rate) by a tax agency. Participants are told that only they know their income and that they may disclose any amount up to the amount of income they have received. They are also told that they may be audited and any income not disclosed will be detected and a fine

imposed. All of the relevant parameters are described in the instructions and are provided on the screen at all times the participants are making their decisions.

The experimental software is highly interactive. The computer screen informs the participants of the base audit probability and penalty. When the participant enters a proposed income disclosure, the screen updates the conditional audit probability. The probability is determined by the formula: $\text{Actual Probability} = \text{Base Probability} + 0.001 (\text{Actual Income} - \text{Disclosed Income})$. The participants are free to experiment with different disclosure decisions until they actually click on the “File Taxes” button. The screen updates and informs the participants of the actual (endogenous) probability of being audited whenever the participant enters an income level to disclose. The screen also informs the participants of the outcome (take home income) that would be added to their balance if they were audited and if they were not audited. While the participants may input different values and observe the prospective results, there is a time limit imposed; participants must click on the “File Taxes” button within 2min and are warned when the time limit is approaching. This simulates the necessity of filing within the legal time limit.

Once all participants have disclosed their income, the audit process begins. While the base audit probability is the same for all participants, the effective audit probabilities differed according to the level of income reported. The computer screen informs the participants of their effective audit probability based on their intended declaration prior to their actually submitting their tax report. After the audit process, the computer screen informs each participant of her individual outcome. If she is audited, she is told the level of the fine imposed and the resulting net income for the period. If she is not audited, she is so informed. The total number of participants audited is announced at the end of each round.

Several treatments are conducted (Table 4). The experiments employ a within subject design. Thus, each participant sees several treatments during a session and the order of the treatments is changed for each session. To ensure that the data sets would encompass a sufficient number of treatments and be comparable, it was decided that the design would involve having each participant participate in three different settings (series A) lasting a total of nine decision rounds (three consecutive rounds in each setting). Since there are four settings in series A, the participants were randomly assigned across these. A second series (series B) of experiments was run in which the only treatment variable was the audit rate, which changed every two rounds, and all participants experienced all the treatments in this series.

Table 4 Experimental design (parameters).

Treatments series A	Audit probability	Fine rate	Expected value of audit	Treatments series B	Audit probability	Fine rate	Expected value of audit
A1	0.10	1.5	0.15	B1	0.10	3.0	0.3
A2	0.30	3.0	0.90	B2	0.20	3.0	0.6
A3	0.10	3.0	0.30	B3	0.30	3.0	0.9
A4	0.30	1.5	0.45	B4	0.40	3.0	1.2

The parameters for each treatment setting are reported in Table 4. The participants received the same income (405 lab dollars) in each round. They were not informed of the number of rounds that a given treatment would be in effect, nor were they informed of the number of treatments they would face during the session. The exchange rate from lab dollars to local currency was announced prior to the start of the experiment. The audit rates reported in Table 6 represent the base audit probability, but the actual audit probability is endogenous since it varies inversely with the amount disclosed, as discussed above. The fine rates represent the multiplier imposed on unpaid taxes if the individual was audited. The expected value of audit is simply the product of the audit probability and fine rate. This single metric is useful for comparing across treatments although it has no behavioral implications.

4.2. Participant pools

Participants (staff and students) from South Africa and Botswana were recruited by personnel at the respective (state) universities.²¹ In total, there were six sessions run in Botswana and South Africa with 99 participants in the former and 88 in the latter. The average age in the Botswana sessions was 25.4 (standard deviation 6.11) and 28.4 (standard deviation 8.61) in South Africa. In all, 33 percent of the South Africa participants were students and 27 percent in Botswana. All participants had prior experience filing their own taxes. The participant pools are not representative of the populations of the respective countries; the pool is younger and better educated than the general population. However, the pools are similar in terms of education and incomes across the two countries, thus facilitating our investigation of the contribution of governance to tax compliance.

For this experimental investigation, the objective was to create a controlled field setting with the properties of a tax-filing problem. This reminds the participants of the naturally occurring setting they face when selecting their tax compliance strategy in the field. Participant earnings were paid in the local currency (Rand and Pula). The payment rate in all sessions is approximately three times the average adult wage in the region. By all casual observations, the participants were highly motivated by the cash payoffs. Analysis of observed behavior across participant pools within country confirmed that these samples could be pooled.

4.3. Hypotheses investigated

The usual “economics of crime” result for tax compliance behavior suggests that by making the evasion gamble less attractive, fewer people will choose to evade. Thus our first hypothesis:

H1. Compliance levels **increase** as the audit probability increases and as the penalty rate increases. This holds for both countries.

The experimental literature suggests that participants will bring to the laboratory their perceptions of the consequences and ethics of tax evasion *when the experimental setting reinforces this through the use of tax language in the experimental instructions*. Since the experimental parameters (tax rate, laboratory income, and enforcement) are the same for both participant pools, the governance background of the two participant

pools constitutes an orthogonal treatment. The central hypothesis is that observed differences in behavior across the pools is due to social or cultural factors. These factors are predicted to lead to systematically different reactions to the *same experimental parameters*. The experiments reported here are intentionally very context intensive. The main hypothesis focuses on differences due to cultural effects. The discussion in Section 3 concerning perceptions of the public sector, the quality of the political institutions, and the level of tax compliance obtained with survey data lead to the following prediction:

H2. The compliance rate will be higher in Botswana than in South Africa, *ceteris paribus*.

Table 5 Summary statistics—average compliance rates (number of participants).

Treatments series A	South Africa	Botswana	Treatments series B	South Africa	Botswana
A1	0.494 (88)	0.617 (99)	B1	0.513 (88)	0.565 (99)
A2	0.618 (88)	0.721 (94)	B2	0.597 (88)	0.660 (99)
A3	0.485 (42)	0.622 (80)	B3	0.637 (88)	0.747 (99)
A4	0.569 (46)	0.418 (20)	B4	0.697 (88)	0.750 (99)

Table 6 Panel Tobit estimation (random effects) with series B data (dependent variable = compliance rate).

Independent variable	Predict	Model 1	Model 2	Model 3
Constant		0.209*** (0.114)	0.264** (0.134)	0.288** (0.128)
Probability audit	+	0.008*** (0.001)	0.008*** (0.001)	0.010*** (0.001)
Age	+	0.009*** (0.003)	0.008** (0.004)	0.008** (0.004)
Occupation (S = 1)	–	0.0613 (0.048)	0.057 (0.053)	0.057 (0.051)
South Africa	–	–0.104** (0.053)	–0.106** (0.053)	–0.109** (0.056)
Lag audit	?		0.025 (0.023)	0.045* (0.024)
Lag number audited	+			–0.019*** (0.005)
Number of observations		1328	1162	1162
Left-censored		94	71	71
Right-censored		246	228	228
Chi-sq		11.50***	18.22***	21.36***
Log likelihood		–701.65	–610.35	–604.79

5. Experimental results

Summary statistics are presented in Table 5. The participants appeared to understand the setting as compliance generally increased with enforcement effort. The data show that there are some clear differences in behavior across countries. This is borne out by the results reported in Tables 6 and 7, which report the econometric estimations of the series B and series A data, respectively. The participants' response to changes in the audit probability indicates that we cannot reject our Hypothesis 1. We can focus on the effects of participant characteristics, country (Botswana is omitted), and individual audit

experience. The dependent variable, Compliance Rate, is censored, so we use a random effects panel Tobit estimation. The results in Table 6 show that individual audit experience affects compliance and that participants in South Africa exhibit lower compliance (statistically significant at the 0.05 level). Thus, we cannot reject our Hypothesis 2, that low quality governance exerts a negative effect on compliance.

In the series A experiments the audit probability and the penalty rate were varied. The intention was that all participants would experience all four treatments. However, the participants were to be randomly assigned an order, and time limits prevented us from completing all possible treatments with all groups. There are substantially fewer observations for series A4. Further, the software reassigned participants to different terminal identifiers for each series, and no consistent means is available to track the participants through all of the treatments they experienced. Since the data constitute a panel within a treatment and we must account for possible serial dependence, we have elected to run treatment-specific models as reported in Table 7. The cost is that we cannot independently obtain coefficients for the enforcement treatment variables. The estimation employed is a random effects panel Tobit. The Age variable has the expected sign while the Occupation dummy variable is not significant in any specification. In three of the four series the coefficient on the country dummy indicates that the compliance rate is statistically lower in South Africa. This is further confirmation, with a different set of treatment parameters, that our hypothesis of lower compliance in South Africa is not rejected by the experimental data.

The observed behavioral differences across the pools might be argued to be due to differences in risk attitudes (e.g., cultural differences toward taking gambles) rather than the institutional features of the fiscal sectors in the countries. The data from

Table 7 Panel Tobit estimation with series a experiment data—treatment specific models (dependent variable =compliance rate).

Independent Variable		Experiment series				
		Predict	Series A1	Series A2	Series A3	Series A4
Constant			0.169 (0.25)	0.487** (0.211)	0.649** (0.356)	0.333 (0.364)
Age	+		0.015** (0.007)	0.009* (0.005)	0.004 (0.009)	0.009 (0.008)
Occupation (S= 1)	–		0.122 (0.113)	0.062 (0.096)	–0.136 (0.154)	–0.185 (0.178)
South Africa	–		–0.0959* (0.062)	–0.1557*** (0.054)	–0.230** (0.091)	–0.059 (0.122)
Number of observations			558	546	366	180
Left censored			54	32	52	41
Right censored			94	98	56	13
Wald Ch-sq			7.58**	10.49**	8.32**	9.05**
Log likelihood			–375.65	–300.80	–239.08	–118.73

the risk attitude experiments allows us to investigate this conjecture as noted above. The risk taking behavior of the pools is not statistically different based on a Chi-square test (at a 0.01 confidence level) and confirmed via a Kolmogorov–Smirnov test. Observed differences in behavior would *not* appear to be due to pool-specific

differences in risk attitudes. While we cannot eliminate all individual factors, the result for risk attitudes strengthens our conjecture that differences are due to cultural factors attributable to differences in governance quality.

6. Concluding comments

Tax compliance (evasion) is a complex decision that is motivated by a variety of factors. The threat of detection and punishment is clearly a factor, and evidence from a variety of sources supports the proposition that increased enforcement leads to increased compliance. However, observed compliance levels are typically higher than warranted by the level of enforcement. This has led to the formation of theories based on exceptional risk aversion (such as prospect theory and rank dependent expected utility). A promising line of inquiry has been the effect of social norms on compliance behavior. There is evidence that these norms are influenced by the tax regime and by the responsiveness of government to the wishes of the citizens. Thus, some cultural differences in compliance behavior are expected, and these differences should be related to tax regimes and government behavior. The results reported in this paper generally support these arguments. We predicted that compliance would be higher in Botswana, and this is confirmed by the results from the survey data and the experimental investigations. An alternative explanation of differences in risk attitudes or a reluctance to engage in gambles is rejected by the data. The two pools exhibit the same attitudes toward risk in a simple context free gamble experiment. This is a useful result; policy makers are able to influence the perception of the public sector much more readily than they can alter the underlying risk behavior of constituents.

The usual caveat regarding the use of laboratory experiments to inform our understanding of behavior in the naturally occurring world is more in force than usual for an investigation such as ours. The cautions of List and Levitt (2007) must be borne in mind here. Does our laboratory setting provide for the necessary degree of “parallelism” to the naturally occurring world that is crucial to generalizing our experimental results beyond the setting of the lab (Plott, 1987)? The experimental setting need not attempt to capture all of the variation in the naturally occurring environment, but it must sufficiently recreate the fundamental elements of the naturally occurring world if the results are to be relevant in policy debates. While our payoffs are relatively small, our experimental setting provides the computations necessary for the decision and a clear link between decisions and rewards, thus reducing the decision costs. More to the point, were our laboratory results not so consistent with the more qualitative results obtained via the surveys, we would be much less willing to argue that we can generalize beyond the lab. As an example of the usefulness of an artefactual field experiment, our paper demonstrates the capability of this investigative tool to address not only important basic behavioral questions but also complex policy questions in ways that are not accessible to other types of economic investigations.

Although one study cannot, in itself, assert a methodological consideration, the consistency of the results we report from our artefactual field experiment and the survey-based results provide support for incorporating the language of the field setting into the laboratory setting as a prerequisite to generalizing the results beyond the lab.

Providing context is important if our objective is to have the participants bring to the lab setting the lessons learned in their life experiences. In this sense, a significant contribution of this paper to the literature on cross-cultural effects on tax compliance is the joint use of survey data on tax morale and an artefactual field experiment framed in tax language that investigates compliance behavior. Our experimental and survey results provide mutual support for the hypothesis that tax morale enhance compliance, for given enforcement effort, and our results demonstrate that the quality of governance has an observable impact on tax compliance precisely because we have implemented our stylized tax setting in the lab rather than using a context-free experimental environment. Our findings further suggest the relevance of models of tax compliance that go beyond the conventional “economics of crime” approach and that capture the role of institutions, in particular how individuals perceive their governments, in explaining why individuals pay taxes. Our results suggest that traditional remedies to induce higher tax compliance will perform much better if accompanied with improvements in governance.

Acknowledgements

We thank our hosts in Pretoria, South Africa, and Gabarone, Botswana, for their hospitality and for making this research possible. We also thank the participants of the Annual Meeting of the Public Choice Society 2004 in Baltimore, especially Stephen Knack, Julian Westerhout and Christopher Westley for their helpful comments. The comments of three anonymous referees and John List were very much appreciated, and the resulting paper is much improved.

References

- Acemoglu, D., Johnson, S., Robinson, A., 2002. An African success story: Botswana. CEPR Discussion Paper 3219.
- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003. Fractionalization. *Journal of Economic Growth* 8, 155–194.
- Allingham, M.G., Sandmo, A., 1972. Income tax evasion: a theoretical analysis. *Journal of Public Economics* 1, 323–338.
- Alm, J., 1988. Uncertain tax policies, individual behavior, and welfare. *American Economic Review* 78, 237–245.
- Alm, J., Jackson, B.R., McKee, M., 1992a. Institutional uncertainty and taxpayer compliance. *American Economic Review* 82, 1018–1026.
- Alm, J., Jackson, B.R., McKee, M., 1993. Fiscal exchange, collective decision institutions, and tax compliance. *Journal of Economic Behavior and Organization* 22, 285–303.
- Alm, J., McClelland, G.H., Schulze, W.D., 1999. Changing the social norm of tax compliance by voting. *Kyklos* 52, 141–171.
- Alm, J., McClelland, G.H., Schulze, W.D., 1992b. Why do people pay taxes? *Journal of Public Economics* 48, 21–38.
- Alm, J., Martinez-Vazquez, J., 2003. Institutions, paradigms, and tax evasion in developing and transition countries. In: Martinez-Vazquez, J., Alm, J. (Eds.), *Public Finance in Developing and Transition Countries*. Edgar Elgar, UK, pp. 146–178.
- Alm, J., McKee, M., 2006. Audit certainty and taxpayer compliance. *National Tax Journal* 59, 801–816.
- Andreoni, J., Erard, B., Feinstein, J., 1998. Tax compliance. *Journal of Economic Literature* 36, 818–860.
- Bernasconi, M., 1998. Tax evasion and orders of risk aversion. *Journal of Public Economics* 67, 123–134.
- Clotfelter, C., 1983. Tax evasion and tax rates. *Review of Economics and Statistics* 65, 363–373.
- Cowell, F.A., 1990. *Cheating the Government: The Economics of Evasion*. MIT Press, Cambridge, MA.
- Deininger, K., Squire, L., 1996. A new data set measuring income inequality. *World Bank Economic Review* 10, 565–591.
- Elster, J., 1989. Social norms and economic theory. *The Journal of Economic Perspectives* 3, 99–117.
- Feinstein, J.S., 1991. An econometric analysis of income tax evasion and its detection. *RAND Journal of Economics* 42, 14–35.

- Forest, A., Sheffrin, S.M., 2002. Complexity and compliance: an empirical investigation. *National Tax Journal* 55, 75–88.
- Galbraith, J.K., Kum, H., 2005. Estimating the inequality of household income: a statistical approach to the creation of a dense and consistent data set. *Review of Income and Wealth* 51, 115–143.
- Gneezy, U., Rustichini, A., 2000. A fine is a price. *Journal of Legal Studies* 29, 1–17.
- Harrison, G.W., List, J.A., 2004. Field experiments. *Journal of Economic Literature* 42, 1009–1055.
- Holt, C., Laury, S., 2002. Risk aversion and incentive effects. *American Economic Review* 95, 1644–1655.
- Index of Economic Freedom, 2005. <http://www.heritage.org/index/>.
- Kaufmann, D., Kraay, A., Mastruzzi, M., 2003. Governance Matters III: Governance Indicators for 1996–2002. World Bank, Washington D.C.
- Kinsey, K., 1992. Deterrence and alienation effects of IRS enforcement: an analysis of survey data. In: Slemrod, J. (Ed.), *Why People Pay Taxes: Tax Compliance and Enforcement*. University of Michigan Press, Ann Arbor, MI, pp. 259–285.
- Knack, S., Keefer, P., 1997. Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics* 112, 1251–1288.
- Knack, S., 1999. Aid dependence and the quality of governance. IRIS Center, World Bank Policy Research Working Paper No. 2396, available at SSRN: <http://ssrn.com/abstract=630769>.
- List, J., Levitt, S., 2007. What do laboratory experiments tell us about the real world? *Journal of Economic Perspectives* 21, 153–174.
- Naylor, R., 1989. Strikes, free riders, and social consensus. *Quarterly Journal of Economics* 104, 771–786.
- Plott, C.R., 1987. Dimensions of parallelism: some policy applications of experimental methods. In: Roth, A.E. (Ed.), *Laboratory Experimentation in Economics: Six Points of View*. Cambridge University Press, New York, NY, pp. 193–219.
- Polity IV, 2002. <http://www.systemicpeace.org/polity/polity4.htm>.
- Roth, A.E., Prasnikar, V., Okuno-Fujiwara, M., Zamir, S., 1991. Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study. *American Economic Review* 81, 1068–1095.
- Schönteich, M., 2000. South Africa's position in Africa's crime rankings. *African Security Review* 9. Available at <http://www.iss.co.za/Pubs/ASR/9No4/Contents.html>, accessed, September 12, 2005.
- Slemrod, J., Blumenthal, M., Christian, C., 2001. Taxpayer response to an increased probability of an audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics* 79, 455–483.
- Smith, K.W., 1992. Reciprocity and fairness: positive incentives for tax compliance. In: Slemrod, J. (Ed.), *Why People Pay Taxes: Tax Compliance and Enforcement*. University of Michigan Press, Ann Arbor, MI, pp. 223–258.

Smith, K.W., Stalans, L.J., 1991. Encouraging tax compliance with positive incentives: a conceptual framework and research directions. *Law and Society Review* 13, 35–53.

Steenbergen, M.R., McGraw, K.H., Scholz, J.T., 1992. Taxpayer adaptation to the 1986 Tax Reform Act: do new tax laws affect the way taxpayers think about taxes? In: Slemrod, J. (Ed.), *Why People Pay Taxes: Tax Compliance and Enforcement*. University of Michigan Press, Ann Arbor, MI, pp. 9–37.

The Star, October 28, 1999. Johannesburg, South Africa, p. 6. Transparency International, 1999. <http://www.transparency.org/>.

Wartick, M.L., Madeo, S.A., Vines, C., 1999. Reward dominance in tax reporting experiments: the role of context. *Journal of the American Taxation Association* 21, 20–31.

Yitzhaki, S., 1974. A note on income tax evasion: a theoretical analysis. *Journal of Public Economics* 3, 201–202.